SAG19: exoplanet imaging signal detection theory and rigorous contrast metrics

Chairs: Dimitri Mawet (Caltech) and Rebecca Jensen-Clem (Caltech)

Team members: John Krist (JPL), Ruslan Belikov (NASA AMES), Steve Bryson (NASA AMES), Jean-Baptiste Ruffio (Stanford), Johan Mazoyer (STScI), Brendan Crill (JPL), Tiffany Glassman (Northrop), Maggie Turnbull (SETI), Angelle Tanner (MSU), T.J. Rodigas (Carnegie DTM), M. Kenworthy (Leiden), Thayne Currie (Subaru), Garreth Ruane (CIT), John Trauger (JPL), Carlos Gomez (ULg), Elodie Choquet (JPL)

As planning for the next generation of high contrast imaging instruments (e.g. WFIRST, HabEx, and LUVOIR, TMT-PFI, EELT-EPICS) matures, and second-generation ground-based extreme adaptive optics facilities (e.g. VLT-SPHERE, Gemini-GPI) are halfway through their large main surveys, it is imperative that the performance of different designs, post-processing routines, observing strategies, and survey results be compared in a consistent, statistically robust framework. SAG19, exoplanet imaging signal detection theory and rigorous contrast metrics, is overarching to all direct imaging instrument, strategies, and methods. The scope of SAG19 is:

1- To go back to the basics of Bayesian Signal Detection Theory (SDT).
Bayesian SDT implies H0:signal absent / H1:signal present hypothesis testing, and invokes well-known concepts such as: the confusion/contingency matrix, false positive (type I error), false negative (type II error), true positive, and true negative fractions, and useful combinations of these quantities such as sensitivity (or completeness) and specificity.

2- To rebuild a solid set of usual definitions used for or in lieu of “contrast” in different contexts, such as astrophysical contrast or ground truth, instrumental contrast used for coronagraph/instrument designs, and the measured on-sky data-driven contrast.

Bayesian, hypothesis testing SDT will automatically force our community to be inclusive of all possible aspects of exoplanet detection, and signal-to-noise ratio (SNR) metrics, including signal-related parameters: planet-star contrast, SED, polarization, variability; instrument parameters: throughput, bandwidth, Strehl ratio/encircled energy, background (sky/thermal, or astrophysical), detector characteristics; noise characteristics as affected by the starlight suppression technique (in a very broad sense): mean intensity, RMS pixel intensities, RMS resolution element (resel, of characteristic size wavelength/telescope diameter) intensities, the probability density function (PDF) computed over pixels, the same PDF computed over resels, their nature and higher order moments, the sample zone and size, outlier management, etc.

3- To identify what we can learn and apply from communities outside our field (e.g. medical imaging). A good example is the widespread use of receiver operating characteristic curve (ROC) and area under the curve (AUC). ROC plots the true positive fraction against the false positive fraction, and is useful to capture the true performance of a given high contrast imaging instrument, or post-processing/detection algorithm. Other formalisms, alternative to the ROC curve, such as the precision-recall curve will also be considered.

4- To define precise contrast computation and ROC curve computation recipes, a new “industry standard”. The goal is to be able to compare results from surveys, instrument and/or telescope designs on a level-playing field.

5- To identify how the new metrics and recipes can be used to define confidence levels for detection (H1) and subsequently error bars for photometric, spectroscopic, astrometric characterization. Ancillary goals: better understanding what limits exoplanet characterization, not just detection. For instance, understanding the limiting precision of extracted planet spectra from algorithms that anneal the planet signal and gaining proper error assessments from spectral extraction. This is particularly important in cases where the prior on the signal/wavelength to be detected is unknown and iterative forward-modeling must be applied.

6- To perform a community data challenge before and after applying our proposed set of standardized SDT rules and recipes, and apply lessons learned.